UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER OF PATENTS AND TRADEMARKS
Washington, D.C. 20231
www.uspto.gov

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
09/736,968	12/13/2000	Peter S. Lu	020054-000611US	8280

20350 7590 11/20/2002

TOWNSEND AND TOWNSEND AND CREW, LLP
TWO EMBARCADERO CENTER
EIGHTH FLOOR
SAN FRANCISCO, CA 94111-3834

EXAMINER
BUNNER, BRIDGET E

ART UNIT	PAPER NUMBER
1647	15

DATE MAILED: 11/20/2002

Please find below and/or attached an Office communication concerning this application or proceeding.

Office Action Summary	Application No.	Applicant(s)	
	09/736,968	LU ET AL.	
	Examiner	Art Unit	
	Bridget E. Bunner	1647	

-- The MAILING DATE of this communication appears on the cover sheet with th corresp ndence address --

Period for Reply

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE 3 MONTH(S) FROM THE MAILING DATE OF THIS COMMUNICATION.

- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
- If the period for reply specified above is less than thirty (30) days, a reply within the statutory minimum of thirty (30) days will be considered timely.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133).
- Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

Status

1)☒ Responsive to communication(s) filed on *04 September 2002* .

2a)☐ This action is **FINAL**. 2b)☒ This action is non-final.

3)☐ Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

Disposition of Claims

4)☒ Claim(s) *1-37* is/are pending in the application.

4a) Of the above claim(s) *5,16-29 and 31-37* is/are withdrawn from consideration.

5)☐ Claim(s) _____ is/are allowed.

6)☒ Claim(s) *1-4,6-15 and 30* is/are rejected.

7)☐ Claim(s) _____ is/are objected to.

8)☒ Claim(s) *1-37* are subject to restriction and/or election requirement.

Application Papers

9)☒ The specification is objected to by the Examiner.

10)☐ The drawing(s) filed on _____ is/are: a)☐ accepted or b)☐ objected to by the Examiner.

Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).

11)☐ The proposed drawing correction filed on _____ is: a)☐ approved b)☐ disapproved by the Examiner.

If approved, corrected drawings are required in reply to this Office action.

12)☐ The oath or declaration is objected to by the Examiner.

Priority under 35 U.S.C. §§ 119 and 120

13)☐ Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).

a)☐ All b)☐ Some * c)☐ None of:

1.☐ Certified copies of the priority documents have been received.

2.☐ Certified copies of the priority documents have been received in Application No. _____ .

3.☐ Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).

* See the attached detailed Office action for a list of the certified copies not received.

14)☒ Acknowledgment is made of a claim for domestic priority under 35 U.S.C. § 119(e) (to a provisional application).

a) ☐ The translation of the foreign language provisional application has been received.

15)☒ Acknowledgment is made of a claim for domestic priority under 35 U.S.C. §§ 120 and/or 121.

Attachment(s)

1) ☒ Notice of References Cited (PTO-892)
2) ☐ Notice of Draftsperson's Patent Drawing Review (PTO-948)
3) ☐ Information Disclosure Statement(s) (PTO-1449) Paper No(s) _____ .
4) ☐ Interview Summary (PTO-413) Paper No(s). _____ .
5) ☐ Notice of Informal Patent Application (PTO-152)
6) ☐ Other: .

DETAILED ACTION

Status of Application, Amendments and/or Claims

The amendment of 12 October 2001 (Paper No. 8) has been entered in full.

Election/Restrictions

Applicant's election with traverse of Applicant's election without traverse of Group A, claims 1-4, 6-15, and 30, drawn to an isolated CLASP-7 polynucleotide, an expression vector, a host cell system, and a method of producing a CLASP-7 polypeptide in Paper No. 14 (04 September 2002) is acknowledged.

Applicant's election with traverse of Group A, claims 1-6, 8-17, and 32, drawn to an isolated CLASP-2 polynucleotide, an expression vector, a host cell system, and a method of producing a CLASP-2 polypeptide in Paper No. 13 (23 August 2002) is acknowledged. The traversal is on the ground(s) that Groups A, F, and G should be rejoined. This is not found persuasive because, as discussed in the previous Office Action (30 July 2002), the polynucleotide of Invention A can be used in materially different processes other than the methods of Groups F and G, such as DNA purification and gene therapy (MPEP § 086.05(h)). Each of groups A, F, and G are unique invention, requiring a unique search of the prior art. Searching all of the inventions in a single patent application would provide an undue search burden on the examiner and the USPTO's resources because of the non-coextensive nature of these searches.

The requirement is still deemed proper and is therefore made FINAL.

Claims 5, 16-29, and 31-37 are withdrawn from further consideration pursuant to 37 CFR 1.142(b), as being drawn to a nonelected group, there being no allowable generic or linking

claim. Applicant timely traversed the restriction (election) requirement in Paper No. 14 (04 September 2002).

It is noted to Applicant that although claims 28 and 29 may recite unintentional errors, Applicant must comply with 37 CFR 1.121 to make claim amendments.

Claims 1-4, 6-15, and 30 are under consideration in the instant application.

Specification

1. The disclosure is objected to because of the following informalities:

1a. Patent applications are referenced throughout the disclosure (pg 2, line 7; pg 111, line 22). The status of the applications must be updated.

1b. The disclosure is objected to because it contains an embedded hyperlink and/or other form of browser-executable code (See page 12, line 17). Applicant is required to delete the embedded hyperlink and/or other form of browser-executable code. See MPEP § 608.01.

1c. The specification is replete with blanks, which are not clear, concise, and exact (see for example, pages 3, 31, 111).

1d. The Brief Description of Drawings refers to Figure 5C. However, Figure 5C is identified as Figure 5B (2 of 2) on the Drawing sheet.

1e. The title of the invention is not descriptive. A new title is required that is clearly indicative of the invention to which the claims are directed.

The following title is suggested: "NUCLEIC ACID MOLECULE ENCODING A CLASP-7 TRANSMEMBRANE PROTEIN".

Appropriate correction is required.

Claim Objections

2. Claims 1 and 30 are objected to because of the following informalities:

2a. Claim 1(c) at line 3 is missing the term "least" before the phrase "25 contiguous residues".

2b. Claim 30 recites non-elected groups.

Appropriate correction is required.

Claim Rejections - 35 USC § 101 and § 112, first paragraph

35 U.S.C. 101 reads as follows:

> Whoever invents or discovers any new and useful process, machine, manufacture, or composition of matter, or any new and useful improvement thereof, may obtain a patent therefore, subject to the conditions and requirements of this title.

The following is a quotation of the first paragraph of 35 U.S.C. 112:

> The specification shall contain a written description of the invention, and of the manner and process of making and using it, in such full, clear, concise, and exact terms as to enable any person skilled in the art to which it pertains, or with which it is most nearly connected, to make and use the same and shall set forth the best mode contemplated by the inventor of carrying out his invention.

3. Claims 1-4, 6-15, and 30 are rejected under 35 U.S.C. 101 because the claimed invention is not supported by either a credible, specific and substantial asserted utility or a well established utility. Novel biological molecules lack well established utility and must undergo extensive experimentation.

Specifically, claims 1-4, 6-15, and 30 are directed to an isolated Cadherin-like asymmetry protein-7 (CLASP-7) polynucleotide wherein the polynucleotide is (a) a polynucleotide that has the sequence of SEQ ID NO: 1, (b) a polynucleotide that hybridizes under stringent hybridization conditions to (a) and encodes a polypeptide having the sequence of SEQ ID NO: 2 or an allelic variant or homologue, (c) a polynucleotide that hybridizes under stringent hybridization conditions to (a) and encodes a polypeptide with at least 25 contiguous residues of the

polypeptide of SEQ ID NO: 2, or (d) a polynucleotide that hybridizes under stringent hybridization conditions to (a) and has at least 12 contiguous bases identical to or exactly complementary to SEQ ID NO: 1. The claims further recite that the polynucleotide encodes a polypeptide having the sequence of SEQ ID NO: 2. The claims recite an isolated CLASP-7 polynucleotide comprising a nucleotide sequence that has at least 90% identity to SEQ ID NO: 1. The claims also recite an expression vector comprising the polynucleotide, a host cell, and a method for producing the polypeptide. Additionally, the claims are directed to an antisense oligonucleotide complementary to a mRNA comprising SEQ ID NO: 1 and an antisense polynucleotide less than about 200 bases in length. The claims recite a pharmaceutical composition comprising a polynucleotide and a pharmaceutically acceptable carrier.

The specification asserts that the CLASP-7 polynucleotide (SEQ ID NO:1) and polypeptide (SEQ ID NO: 2) of the present invention are involved in a variety of cellular processes, particularly related to immune function, T cell activation, regulation of T cell and B cell interactions, and in the organization, establishment, and maintenance of the "immunological synapse" (including signal transduction, cytoskeletal interactions, and membrane organization) (pg 18, lines 6-13). The specification also discloses that the CLASP-7 protein is believed to be a component of the lymphocyte organelle called the "immune gateway" that creates a docking site or portal for cell-cell contact during antigen presentation (pg 18, lines 14-17). However, the instant specification does not teach any significance or functional characteristics of the CLASP-7 polynucleotide (SEQ ID NO: 1) or polypeptide (SEQ ID NO: 2). The specification also does not disclose any methods or working examples that indicate the polynucleotide and polypeptide of the instant invention are involved in any of the abovementioned activities. Since significant

further research would be required of the skilled artisan to determine how the claimed polypeptide is involved with the above-mentioned activities, the asserted utilities are not substantial. Since the utility is not presented in mature form and significant further research is required, the utility is not substantial. The specification asserts the following as patentable utilities for the claimed putative polynucleotide (SEQ ID NO: 1):

1) to detect the expression of CLASP-7 in cells (pg 41, lines 32-33; pg 42; pg 48-50; pg 76-77)

2) in the diagnosis of a disorder or disease resulting from aberrant expression of CLASP-7 (pg 42, lines 1-9)

3) as hybridization probes for cDNA and genomic DNA (pg 42, lines 20-33; pg 44; pg 45, lines 1-25)

4) as primers for a nucleic acid amplification pg 43, lines 20-33; pg 44; pg 45, lines 1-25)

5) to treat, detect, or modulate immune system disorders, hematopoietic cell disorders, allergic reactions, organ rejection or graft-versus-host disease, inflammation, infectious agents (pg 44, lines 26-34; 45-47)

6) to engineer hammerhead motif ribozyme molecules (pg 54, lines 33-34; pg 55, lines 1-13)

7) for gene therapy (pg 55-60)

8) to construct a transgenic animal (pg 61-62)

9) in chromosome mapping (pg 62, lines 25-34; pg 63-64)

10) to screen CLASP-7 agonists and antagonists (pg 40, lines 32-33; pg 41, lines 1-2)

Each of these shall addressed in turn.

1) to detect the expression of CLASP-7 in cells. This asserted utility is credible, but not specific or substantial. The specification does not disclose a specific target sequence. The specification does not disclose the cell types that express CLASP-7. Significant further

experimentation would be required of the skilled artisan to identify cells with CLASP-7. Since this asserted utility is also not present in mature form so that it could be readily used in real world sense, the asserted utility is not substantial.

2) in the diagnosis of a disorder or disease resulting from aberrant expression of CLASP-7. This asserted utility is credible but not specific or substantial. The specification does not disclose disorders associated with a mutated, deleted, or translocated CLASP-7 gene. Significant further experimentation would be required of the skilled artisan to identify individuals with such a disease. Since this asserted utility is also not present in mature form so that it could be readily used in real world sense, the asserted utility is not substantial.

3) as hybridization probes for cDNA and genomic DNA. This asserted utility is credible but not substantial or specific. Hybridization probes can be designed from any polynucleotide sequence. Further, the specification does not disclose specific cDNA, DNA, or RNA targets. Since this asserted utility is also not present in mature form, so that it could be readily used in a real world sense, the asserted utility is not substantial.

4) as primers for a nucleic acid amplification. This asserted utility is credible but not substantial or specific. Primers can be designed from any polynucleotide sequence. Further, the specification does not disclose a specific DNA target. Since this asserted utility is also not present in mature form, so that it could be readily used in a real world sense, the asserted utility is not substantial.

5) to treat, detect, or modulate immune system disorders, hematopoietic cell disorders, allergic reactions, organ rejection or graft-versus-host disease, inflammation, infectious agents. This asserted utility is credible but not specific or substantial. The specification does not

disclose disorders associated with a mutated, deleted, or translocated CLASP-7 gene (SEQ ID NO: 1). The specification does not disclose which disorders are associated with altered levels of the CLASP-7 gene. Significant further experimentation would be required of the skilled artisan to identify individuals with such a disease. Since this asserted utility is also not present in mature form, so that it could be readily used in a real world sense, the asserted utility is not substantial.

6) to engineer hammerhead motif ribozyme molecules. This asserted utility is credible but not specific or substantial. Ribozymes can be designed from any DNA/RNA sequence. Additionally, the specification does not disclose a specific DNA/RNA target. Since this asserted utility is also not present in mature form, so that it could be readily used in a real world sense, the asserted utility is not substantial.

7) for gene therapy. This asserted utility is not credible, specific or substantial. The specification does not disclose diseases associated with a mutated, deleted, or translocated CLASP-7 gene of SEQ ID NO: 1. Significant further experimentation would be required of the skilled artisan to identify individuals with such a disease. Since this asserted utility is also not presented in mature form, so that it could be readily used in a real world sense, the asserted utility is not substantial.

8) to construct a transgenic animal. This asserted utility is credible but not specific or substantial. The specification does not disclose diseases associated with a mutated, deleted, or translocated CLASP-7 gene (SEQ ID NO: 1). Significant further experimentation would be required of the skilled artisan to identify such a disease. The specification discloses nothing about whether the gene will be "knocked in" or "knocked out" or what specific tissues and cells

are being targeted. Since this asserted utility is also not present in mature form, so that it could be readily used in a real world sense, the asserted utility is not substantial.

9) in chromosome mapping. This asserted utility is credible but not substantial or specific. Such assays can be performed with any polynucleotide. Further, the specification does not disclose a specific DNA target. Since this asserted utility is also not present in mature form, so that it could be readily used in a real world sense, the asserted utility is not substantial.

10) to screen CLASP-7 agonists and antagonists. This asserted utility is credible but not specific or substantial. Such assays can be performed with any polynucleotide. Nothing is disclosed about how the polynucleotide is affected by the compounds. Additionally, the specification discloses nothing specific or substantial for the CLASP-7 agonists and antagonists screened in this method. Since this asserted utility is also not presented in mature form, so that it could be readily used in a real world sense, the asserted utility is not substantial.

4. Claims 1-4, 6-15, and 30 are also rejected under 35 U.S.C. 112, first paragraph. Specifically, since the claimed invention is not supported by either a specific and substantial asserted utility or a well established utility for the reasons set forth above, one skilled in the art clearly would not know how to use the claimed invention.

5. Furthermore, claims 1-4, 6-11, 14-15, and 30 recite a CLASP-7 polynucleotide wherein the polynucleotide is a polynucleotide that hybridizes under stringent hybridization conditions to (a) and encodes a polypeptide having the sequence of SEQ ID NO: 2 or an allelic variant or homologue, a polynucleotide that hybridizes under stringent hybridization conditions to (a) and encodes a polypeptide with at least 25 contiguous residues of the polypeptide of SEQ ID NO: 2, or (d) a polynucleotide that hybridizes under stringent hybridization conditions to (a) and has at

least 12 contiguous bases identical to or exactly complementary to SEQ ID NO: 1. The claims also recite an isolated CLASP-7 polynucleotide comprising a nucleotide sequence that has at least 90% identity to SEQ ID NO: 1. Additionally, the claims are directed to an antisense polynucleotide less than about 200 bases in length.

The specification discloses that "the CLASP-7 variants of the invention can contain alterations in the coding regions, non-coding regions, or both" (pg 38, lines 19-20). The specification teaches that known methods of protein engineering and recombinant DNA technology can generate variants to improve or alter the characteristics of the CLASP-7 polypeptides (pg 39, lines 9-11). However, the specification does not teach any allelic variants or homologs of the CLASP-7 polynucleotide or polypeptide. The specification does not disclose (i) a polynucleotide that encodes a polypeptide with at least 25 contiguous residues of the polypeptide of SEQ ID NO: 2, (ii) a polynucleotide that has at least 12 contiguous bases identical to or exactly complementary to SEQ ID NO: 1, or (iii) an antisense polynucleotide less than about 200 bases in length. The specification also does not teach a nucleic acid sequence with 90% sequence identity to the nucleotide sequence of SEQ ID NO: 1. Furthermore, regarding allelic variants, it is noted that such are recognized in the art as variant genes which map to the same locus on the chromosome (See Lewin, Genes II, 1985, pg 681). The specification does not disclose the chromosomal location of any of CLASP-7 gene characterized by the inventors. Additionally, the specification does not teach functional or structural characteristics of any polynucleotide variants in the context of a cell or organism.

Generally, the art acknowledges that function cannot be predicted based solely on structural similarity to a protein found in the sequence databases. For example, Skolnick et al.

(2000, Trends in Biotech. 18:34-39) state that knowing the protein structure by itself is insufficient to annotate a number of functional classes, and is also insufficient for annotating the specific details of protein function (see Box 2, p. 36). Similarly, Bork (2000, Genome Research 10:398-400) states that the error rate of functional annotations in the sequence database is considerable, making it even more difficult to infer correct function from a structural comparison of a new sequence with a sequence database (see especially p. 399). Such concerns are also echoed by Doerks et al. (1998, Trends in Genetics 14:248-250) who state that (1) functional information is only partially annotated in the database, ignoring multi functionality, resulting in underpredictions of functionality of a new protein and (2) overpredictions of functionality occur because structural similarity often does not necessarily coincide with functional similarity. Smith et al. (1997, Nature Biotechnology 15:1222-1223) remark that there are numerous cases in which proteins having very different functions share structural similarity due to evolution from a common ancestral gene. Brenner (1999, Trends in Genetics 15:132-133) argues that accurate inference of function from homology must be a difficult problem since, assuming there are only about 1000 major gene superfamilies in nature, then most homologs must have different molecular and cellular functions. Finally, Bork et al. (1996, Trends in Genetics 12:425-427) add that the software robots that assign functions to new proteins often assign a function to a whole new protein based on structural similarity of a small domain of the new protein to a small domain of a known protein. Such questionable interpretations are written into the sequence database and are then considered facts.

Therefore, based on the discussions above concerning the specific examples of structurally similar proteins that have different functions, along with the art's recognition that

one cannot rely upon structural similarity alone to determine functionality, the specification fails to teach the skilled artisan how to use the claimed polynucleotide to make biologically active CLASP-7 without resorting to undue experimentation to determine what the specific biological activities of the polypeptide are.

The specification does not teach the skilled artisan how to use the claimed polynucleotides encoding CLASP-7 for purposes unrelated to the asserted biological activity. For example, there is no disclosure of particular disease states correlating to an alteration in levels or forms of the polypeptide such that the claimed polynucleotide encoding CLASP-7 could be used as a diagnostic tool. Therefore, the skilled artisan is not provided with sufficient guidance to use the claimed polynucleotides for any purpose.

Further, the problem of predicting protein and DNA structure from sequence data and in turn utilizing predicted structural determinations to ascertain functional aspects of the protein and DNA is extremely complex. For example, while it is known that many amino acid substitutions are generally possible in any given protein the positions within the protein's sequence where such amino acid substitutions can be made with a reasonable expectation of success are limited. Certain positions in the sequence are critical to the protein's structure/function relationship, e.g. such as various sites or regions directly involved in binding, activity and in providing the correct three-dimensional spatial orientation of binding and active sites. These or other regions may also be critical determinants of antigenicity. These regions can tolerate only relatively conservative substitutions or no substitutions (see Wells, 1990, Biochemistry 29:8509-8517; Ngo et al., 1994, The Protein Folding Problem and Tertiary Structure Prediction, pp. 492-495). However, Applicant has provided little or no guidance beyond the mere presentation of sequence data to

enable one of ordinary skill in the art to determine, without undue experimentation, the positions in the protein and DNA which are tolerant to change and the nature and extent of changes that can be made in these positions. The art recognizes that function cannot be predicted from structure alone (Bork, 2000, Genome Research 10:398-400; Skolnick et al., 2000, Trends in Biotech. 18(1):34-39, especially p. 36 at Box 2; Doerks et al., 1998, Trends in Genetics 14:248-250; Smith et al., 1997, Nature Biotechnology 15:1222-1223; Brenner, 1999, Trends in Genetics 15:132-133; Bork et al., 1996, Trends in Genetics 12:425-427).

Furthermore, claim 30 is directed to a pharmaceutical composition comprising a CLASP-7 polynucleotide and a pharmaceutically acceptable carrier. The specification teaches a composition comprising an isolated CLASP-7 polynucleotide consisting of the nucleic acid sequence of SEQ ID NO: 1. The specification does not teach how to use a CLASP-7 "pharmaceutical" composition and a "pharmaceutically acceptable carrier" without undue experimentation for the treatment of a disease in an animal. The specification lists disorders to be treated (pg 44-47), but there are no working examples directed to a particular disorder in an animal or administration of the CLASP-7 polynucleotide consisting of the nucleic acid sequence of SEQ ID NO: 1 to an animal for treatment. (Note, this issue could be overcome by deleting the terms "pharmaceutical" and "pharmaceutically acceptable" from the claims.)

Due to the large quantity of experimentation necessary to determine an activity or property of the disclosed polypeptide such that it can be determined how to use the claimed polynucleotides encoding CLASP-7, to generate the infinite number of derivatives recited in the claims and possibly screen same for activity, and to determine the quantity of CLASP-7 polynucleotide to be administered, the most effective administration route, and the duration of

the treatment, the lack of direction/guidance presented in the specification regarding same and the lack of direction/guidance presented in the specification regarding the chromosomal locus of the CLASP-7 gene disclosed in the specification, the absence of working examples directed to same, the complex nature of the invention, the state of the prior art establishing that biological activity cannot be predicted based on structural similarity and the unpredictability of the effects of mutation on protein structure and function, the unpredictability of the effects of the CLASP-7 polynucleotide *in vivo* and the unpredictable nature of the locus for any isolated gene, and the breadth of the claims which fail to recite particular biological activities and also embrace a broad class of structural fragments and variants and which recite any allelic variant, undue experimentation would be required of the skilled artisan to make and/or use the claimed invention in its full scope.

6. Claims 1-4, 6-11, 14-15, and 30 are rejected under 35 U.S.C. 112, first paragraph, as containing subject matter which was not described in the specification in such a way as to reasonably convey to one skilled in the relevant art that the inventor(s), at the time the application was filed, had possession of the claimed invention.

Specifically, claims 1-4, 6-11, 14-15, and 30 are directed a CLASP-7 polynucleotide wherein the polynucleotide is a polynucleotide that hybridizes under stringent hybridization conditions to (a) and encodes a polypeptide having the sequence of SEQ ID NO: 2 or an allelic variant or homologue, a polynucleotide that hybridizes under stringent hybridization conditions to (a) and encodes a polypeptide with at least 25 contiguous residues of the polypeptide of SEQ ID NO: 2, or (d) a polynucleotide that hybridizes under stringent hybridization conditions to (a)

and has at least 12 contiguous bases identical to or exactly complementary to SEQ ID NO: 1. The claims also recite an isolated CLASP-7 polynucleotide comprising a nucleotide sequence that has at least 90% identity to SEQ ID NO: 1. Additionally, the claims are directed to an antisense polynucleotide less than about 200 bases in length.

The specification teaches human a CLASP-7 polynucleotide and polypeptide (SEQ ID NO: 1 and SEQ ID NO: 2, respectively). The specification also discloses that "the CLASP-7 variants of the invention can contain alterations in the coding regions, non-coding regions, or both" (pg 38, lines 19-20). The specification teaches that known methods of protein engineering and recombinant DNA technology can generate variants to improve or alter the characteristics of the CLASP-7 polypeptides (pg 39, lines 9-11). However, the specification does not teach functional or structural characteristics of the polynucleotide variants in the context of a cell or organism. The description of one CLASP-7 polynucleotide species (SEQ ID NO: 1) and one CLASP-7 polypeptide species (SEQ ID NO: 2) is not adequate written description of an entire genus of functionally equivalent polynucleotides and polypeptides which incorporate all variants and fragments (i) with at least 90% sequence identity to the human CLASP-7 polynucleotide comprising SEQ ID NO: 1. The description of one CLASP-7 polynucleotide species and one CLASP-7 polypeptide species is also not adequate written description of an entire genus of functionally equivalent polynucleotides and polypeptides which incorporate all variants and fragments (ii) that encode a polypeptide with at least 25 contiguous residues, (iii) have at least 12 bases identical to or exactly complementary to SEQ ID NO: 1, and (iv) have antisense polynucleotide less than 200 bases in length.

Vas-Cath Inc. v. Mahurkar, 19USPQ2d 1111, clearly states that "applicant must convey with reasonable clarity to those skilled in the art that, as of the filing date sought, he or she was in possession *of the invention.* The invention is, for purposes of the 'written description' inquiry, *whatever is now claimed"* (See page 1117). The specification does not "clearly allow persons of ordinary skill in the art to recognize that [he or she] invented what is claimed" (See *Vas-Cath* at page 1116).

With the exception of the sequences referred to above, the skilled artisan cannot envision the detailed chemical structure of the encompassed polynucleotides, and therefore conception is not achieved until reduction to practice has occurred, regardless of the complexity or simplicity of the method of isolation. Adequate written description requires more than a mere statement that it is part of the invention and reference to a potential method of isolating it. The nucleic acid itself is required. See *Fiers v. Revel*, 25 USPQ2d 1601 at 1606 (CAFC 1993) and *Amgen Inc. v. Chugai Pharmaceutical Co. Ltd.*, 18 USPQ2d 1016.

One cannot describe what one has not conceived. See *Fiddes v. Baird*, 30 USPQ2d 1481 at 1483. In *Fiddes*, claims directed to mammalian FGF's were found to be unpatentable due to lack of written description for that broad class. The specification provided only the bovine sequence.

Therefore, only an isolated CLASP-7 polynucleotide that has the nucleotide sequence of SEQ ID NO:1 or a CLASP-7 polynucleotide that encodes a polypeptide having the sequence of SEQ ID NO: 2, but not the full breadth of the claim meets the written description provision of 35 U.S.C. §112, first paragraph. Applicant is reminded that *Vas-Cath* makes clear that the written

description provision of 35 U.S.C. §112 is severable from its enablement provision (see page 1115).

7. Claim 3 is rejected under 35 U.S.C. § 112, first paragraph, as containing subject matter which was not described in the specification in such a way as to reasonably convey to one skilled in the relevant art that the inventor(s), at the time the application was filed, had possession of the claimed invention. The invention appears to employ novel nucleic acid molecules (i.e., clones AVC-PD23, AVC-PD24). Since the nucleic acid molecules are essential to the claimed invention they must be obtainable by a repeatable method set forth in the specification or otherwise readily available to the public. If the nucleic acid molecules are not so obtainable or available, the requirements of 35 U.S.C. § 112 may be satisfied by a deposit of the nucleic acid molecules. The specification does not disclose a repeatable process to obtain the nucleic acid molecules and it is not apparent if the nucleic acid molecules are readily available to the public. It is noted that Applicant has deposited the nucleic acid molecules (pg. 106 of the specification), but there is no indication in the specification as to public availability. If the deposit is made under the Budapest Treaty, then an affidavit or declaration by Applicant, or a statement by an attorney of record over his or her signature and registration number, stating that the specific nucleic acid molecules have been deposited under the Budapest Treaty and that the nucleic acid molecules will be irrevocably and without restriction or condition released to the public upon the issuance of a patent, would satisfy the deposit requirement made herein. If the deposit has not been made under the Budapest Treaty, then in order to certify that the deposit meets the criteria set forth in 37 C.F.R. §§ 1.801-1.809, Applicant may provide assurance of compliance by an

affidavit or declaration, or by a statement by an attorney of record over his or her signature and registration number, showing that:

(a) during the pendency of this application, access to the invention will be afforded to the Commissioner upon request;

(b) all restrictions upon availability to the public will be irrevocably removed upon granting of the patent;

(c) the deposit will be maintained in a public depository for a period of 30 years or 5 years after the last request or for the effective life of the patent, whichever is longer;

(d) a test of the viability of the biological material at the time of deposit will be made (see 37 C.F.R. § 1.807); and

(e) the deposit will be replaced if it should ever become inviable.

Applicant's attention is directed to M.P.E.P. §2400 in general, and specifically to §2411.05, as well as to 37 C.F.R. § 1.809(d), wherein it is set forth that "the specification shall contain the accession number for the deposit, the date of the deposit, the name and address of the depository, and a description of the deposited material sufficient to specifically identify it and to permit examination." The specification should be amended to include such, however, Applicant is cautioned to avoid the entry of new matter into the specification by adding any other information.

35 USC § 112, second paragraph

The following is a quotation of the second paragraph of 35 U.S.C. 112:

> The specification shall conclude with one or more claims particularly pointing out and distinctly claiming the subject matter which the applicant regards as his invention.

8. Claims 1-4, 6-15, and 30 are rejected under 35 U.S.C. 112, second paragraph, as being indefinite for failing to particularly point out and distinctly claim the subject matter which applicant regards as the invention.

9. Regarding claims 1-4, 6-15, and 30, the acronym "CLASP-7" renders the claims vague and indefinite. Abbreviations should be spelled out in all independent claims for clarity.

10. Stringency is relative, and the art does not recognize a single set of conditions as stringent. The specification also does not provide an unambiguous definition for the term. In the absence of a recitation of clear hybridization conditions (e.g., "hybridizes at wash conditions of **A** X SSC and **B** % SDS at **C**°C"), claims 1-4, 6-11, 14-15, and 30 fail to define the metes and bounds of the varying structures of polynucleotides recited in the claimed methods.

Claim Rejections - 35 USC § 102

The following is a quotation of the appropriate paragraphs of 35 U.S.C. 102 that form the basis for the rejections under this section made in this Office action:

> A person shall be entitled to a patent unless –
>
> (a) the invention was known or used by others in this country, or patented or described in a printed publication in this or a foreign country, before the invention thereof by the applicant for a patent.

11. Claims 2, 4, and 12-13 rejected under 35 U.S.C. 102(a) as being anticipated by Nagase et al. (Accession No. AB037816, GenEmbl database, 14 March 2000).

Nagase et al. teach an isolated nucleic acid sequence encoding CLASP-7 of SEQ ID NO: 2 wherein the nucleic acid sequence is 100% identical to the nucleotide sequence of SEQ ID NO: 1 of the instant application (See sequence alignments attached to this Office Action as Appendix A and B). Nagase et al. also teach an antisense oligonucleotide complementary to a messenger RNA comprising SEQ ID NO: 1 and encoding CLASP-7. Please note that the claims of the instant application claim priority under 35 U.S.C. 119(e) to 13 different provisional applications. However, Applicant does not disclose the complete nucleotide sequence of SEQ ID NO: 1 or the complete amino acid sequence of SEQ ID NO: 2 of the instant application in any of the

aforementioned provisional applications. Therefore, the priority date of the claims that recite the nucleotide sequence (SEQ ID NO: 1) and amino acid sequence (SEQ ID NO: 2) of the instant application is determined to be 12/13/2000.

Claim Rejections - 35 USC § 103

The following is a quotation of 35 U.S.C. 103(a) which forms the basis for all obviousness rejections set forth in this Office action:

> (a) A patent may not be obtained though the invention is not identically disclosed or described as set forth in section 102 of this title, if the differences between the subject matter sought to be patented and the prior art are such that the subject matter as a whole would have been obvious at the time the invention was made to a person having ordinary skill in the art to which said subject matter pertains. Patentability shall not be negatived by the manner in which the invention was made.

This application currently names joint inventors. In considering patentability of the claims under 35 U.S.C. 103(a), the examiner presumes that the subject matter of the various claims was commonly owned at the time any inventions covered therein were made absent any evidence to the contrary. Applicant is advised of the obligation under 37 CFR 1.56 to point out the inventor and invention dates of each claim that was not commonly owned at the time a later invention was made in order for the examiner to consider the applicability of 35 U.S.C. 103(c) and potential 35 U.S.C. 102(e), (f) or (g) prior art under 35 U.S.C. 103(a).

12. Claims 1, 6-11, 14-15, and 30 are rejected under 35 U.S.C. 103(a) as being unpatentable over NCI-CGAP (Accession No. AI198543; 10 November 1998; EST database) in view of Sibson et al. (WO 94/01548).

NCI-CGAP teaches an isolated that has at least 12 contiguous bases identical to or exactly complementary to SEQ ID NO: 1 of the instant application and an antisense polynucleotide less than 200 bases in length (See sequence alignment attached to this Office

Action as Appendix C; see nucleotides 607-3 of NCI-CGAP et al.; see nucleotides 5768-6372 of SEQ ID NO: 1 of the instant application, for example).

NCI-CGAP does not teach expression vectors, host cells, or a method of producing a polypeptide.

Sibson et al. discloses that it is generally useful to place a desired cDNA sequence into an expression vector, host cell, and express the encoded protein (see pages 8-13).

It would have been obvious to the person of ordinary skill in the art at the time the invention was made to use NCI-CGAP's DNA and the expression vector, host cell, and method of expressing and then isolating the encoded polypeptide as taught by Sibson et al. in view of Sibson's suggestion that it would be desirable to do so, as cited above.

Conclusion

The art made of record and not relied upon is considered pertinent to applicant's disclosure:

Nagase et al. DNA Research 7(1) : 65-73, 2000.

Any inquiry concerning this communication or earlier communications from the examiner should be directed to Bridget E. Bunner whose telephone number is (703) 305-7148. The examiner can normally be reached on 8:30-5:30 M-F.

If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, Gary Kunz can be reached on (703) 308-4623. The fax phone numbers for the organization where this application or proceeding is assigned are (703) 872-9306 for regular communications and (703) 872-9307 for After Final communications.

Any inquiry of a general nature or relating to the status of this application or proceeding should be directed to the receptionist whose telephone number is (703) 872-9305.

Elizabeth C. Bunner

BEB
Art Unit 1647
November 6, 2002

ELIZABETH [illegible]
PRIMARY EXAMINER

Appendix A


```
RESULT   2
AB037816
LOCUS       AB037816               4886 bp    mRNA    linear   PRI 14-MAR-2000
DEFINITION  Homo sapiens mRNA for KIAA1395 protein, partial cds.
ACCESSION   AB037816
VERSION     AB037816.1  GI:7243170
KEYWORDS    .
SOURCE      Homo sapiens brain cDNA to mRNA, clone_lib:pBluescriptII SK plus
            clone:hj07927.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (sites)
  AUTHORS   Nagase,T., Kikuno,R., Ishikawa,K.I., Hirosawa,M. and Ohara,O.
  TITLE     Prediction of the coding sequences of unidentified human genes.
            XVI. The complete sequences of 150 new cDNA clones from brain which
            code for large proteins in vitro
  JOURNAL   DNA Res. 7 (1), 65-73 (2000)
  MEDLINE   20181126
REFERENCE   2  (bases 1 to 4886)
  AUTHORS   Ohara,O., Nagase,T. and Kikuno,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (31-JAN-2000) Osamu Ohara, Kazusa DNA Research Institute,
            Laboratory of DNA Technology; 1532-3 Yana, Kisarazu, Chiba
            292-0812, Japan (E-mail:cdnainfo@kazusa.or.jp,
            URL:http://www.kazusa.or.jp/huge/, Tel:+81-438-52-3913,
            Fax:+81-438-52-3914)
FEATURES             Location/Qualifiers
     source          1. .4886
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="hj07927"
                     /tissue_type="brain"
                     /clone_lib="pBluescriptII SK plus"
     gene            1. .4884
                     /gene="KIAA1395"
     CDS             <1. .>4884
                     /gene="KIAA1395"
                     /note="Start codon is not identified."
                     /codon_start=1
                     /product="KIAA1395 protein"
                     /protein_id="BAA92633.1"
                     /db_xref="GI:7243171"
                     /translation="DARTMAASERRAFAHKINRTVAAEVRKQVSRERSGSPHSSRRCS
                     SSLGVPLTEVVEPLDFEDVLLSRPPDAEPGPLRDLVEFPADDLELLLQPRECRTTEPG
                     IPKDEKLDAQVRAAVEMYIEDWVIVHRRYQYLSAAYSPVTTDTQRERQKGLPRQVFEQ
                     DASGDERSGPEDSNDSRRGSGSPEDTPRSSGASSIFDLRNLAADSLLPSLLERAAPED
                     VDRRNETLRRQHRPPALLTLYPAPDEDEAVERCSRPEPPREHFGQRILVKCLSLKFEI
                     EIEPIFGILALYDVREKKKISENFYFDLNSDSMKGLLRAHGTHPAISTLARSAIFSVT
                     YPSPDIFLVIKLEKVLQQGDISECCEPYMVLKEVDTAKNKEKLEKLRLAAEQFCTRLG
                     RYRMPFAWTAVHLANIVSSAGQLDRDSDSEGERRPAWTDRRRRGPQDRASSGDDACSF
                     SGFRPATLTVTNFFKQEAERLSDEDLFKFLADMRRPSSLLRRLRPVTAQLKIDISPAP
                     ENPHFCLSPELLHIKPYPDPRGRPTKEILEFPAREVYAPHTSYRNLLYVYPHSLNFSS
                     RQGSVRNLAVRVQYMTGEDPSQALPVIFGKSSCSEFTREAFTPVVYHNKSPEFYEEFK
                     LHLPACVTENHHLLFTFYHVSCQPRPGTALETPVGFTWIPLLQHRRLRTGPFCLPVSV
                     DQPPPSYSVLTPDVALPGMRWVDGHKGVFSVELTAVSSVHPQDPYLDKFFTLVHVLEE
                     GAFPFRLKDTVLSEGNVEQELRASLAALRLASPEPLVAFSHHVLDKLVRLVIRPPIIS
                     GQIVNLGRGAFEAMAHVVSLVHRSLEAAQDARGHCPQLAAYVHYAFRLPGTEPSLPDG
                     APPVTVQAATLARGSGRPASLYLARSKSISSSNPDLAVAPGSVDDEVSRILASKLLHE
                     ELALQWVVSSSAVREAILQHAWFFFQLMVKSMALHLLLGQRLDTPRKLRFPGRFLDDI
                     TALVGSVGLEVITRVHKDVELAEHLNASLAFFLSDLLSLVDRGFVFSLVRAHYKQVAT
                     RLQSSPNPAALLTLRMEFTRILCSHEHYVTLNLPCCPLSPPASPSPSVSSTTSQSSTF
                     SSQAPDPKVTSMFELSGPFRQQHFLAGLLLTELALALEPEAEGAFLLHKKAISAVHSL
                     LCGHDTDPRYAEATVKARVAELYLPLLSIARDTLPRLHDFAEGPGQRSRLASMLDSDT
                     EGEGDIAGTINPSVAMAIAGGPLAPGSRASISQGPPTASRAGCALSAESSRTLLACVL
                     WVLKNTEPALLQRWATDLTLPQLGRLLDLLYLCLAAFEYKGKKAFERINSLTFKKSLD
                     MKARLEEAILGTIGARQEMVRRSRERSPFGNPENVRWRKSVTHWKQTSDRVDKTKDEM
                     EHEALVEGNLATEASLVVLDTLEIIVQTVMLSEARESVLGAVLKVVLYSLGSAQSALF
                     LQHGLATQRALVSKFPELLFEEDTELCADLCLRLLRHCGSRISTIRTHASASLYLLMR
                     QNFEIGHNFARVKMQVTMSLSSLVGTTQNFSEEHLRRSLKTILTYAEEDMGLRDSTFA
                     EQVQDLMFNLHMILTDTVKMKEHQEDPEMLIDLMYRIARGYQGSPDLRLTWLQNMAGK
                     HAELGNHAEAAQCMVHAA"
BASE COUNT      941 a   1605 c   1425 g    915 t
ORIGIN
```

```
  Query Match             76.6%;  Score 4882.8;  DB 9;  Length 4886;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches 4884;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 gacgcgaggaccatggctgcctccgagcgccgcgccttcgcgcacaagatcaacaggacg 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 GACGCGAGGACCATGGCTGCCTCCGAGCGCCGCGCCTTCGCGCACAAGATCAACAGGACG 60

Qy      61 gtggccgcagaggtgcggaagcaggtgtcccgggaacgcagtggctccccccactccagc 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 GTGGCCGCAGAGGTGCGGAAGCAGGTGTCCCGGGAACGCAGTGGCTCCCCCCACTCCAGC 120

Qy     121 aggcgctgcagcagctccctgggggtcccactgactgaagttgtcgagcccctggacttt 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 AGGCGCTGCAGCAGCTCCCTGGGGGTCCCACTGACTGAAGTTGTCGAGCCCCTGGACTTT 180

Qy     181 gaggatgtacttctgagccggccaccagatgctgagcccgggccccctcagggacctggta 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 GAGGATGTACTTCTGAGCCGGCCACCAGATGCTGAGCCCGGGCCCCTCAGGGACCTGGTA 240

Qy     241 gaattcccagctgatgacttggagctgctgctgcagccccgggaatgccggaccacggag 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 GAATTCCCAGCTGATGACTTGGAGCTGCTGCTGCAGCCCCGGGAATGCCGGACCACGGAG 300

Qy     301 cccgggatccccaaggatgaaaaactggatgcccaggtgagggccgcggtggagatgtat 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 CCCGGGATCCCCAAGGATGAAAAACTGGATGCCCAGGTGAGGGCCGCGGTGGAGATGTAT 360

Qy     361 attgaggactgggtcattgtccacagaaggtatcagtacctgagtgcagcatacagcccc 420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     361 ATTGAGGACTGGGTCATTGTCCACAGAAGGTATCAGTACCTGAGTGCAGCATACAGCCCC 420
```

```
Qy    421 gtcaccacagacacacagcgggagcgacagaagggcctccccgccaggtctttgagcag 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 GTCACCACAGACACACAGCGGGAGCGACAGAAGGGCCTCCCCCGCCAGGTCTTTGAGCAG 480

Qy    481 gatgcttctggagacgagaggtccggccctgaggactcgaatgactcccggcgtggctcg 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 GATGCTTCTGGAGACGAGAGGTCCGGCCCTGAGGACTCGAATGACTCCCGGCGTGGCTCG 540

Qy    541 ggctccccggaagacacccctcgaagcagtggtgcctctagcatcttcgacctgaggaac 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 GGCTCCCCGGAAGACACCCCTCGAAGCAGTGGTGCCTCTAGCATCTTCGACCTGAGGAAC 600

Qy    601 ctggcagctgactcattgctgccctctctgctagagcgggcggccccagaagatgtggac 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 CTGGCAGCTGACTCATTGCTGCCCTCTCTGCTAGAGCGGGCGGCCCCAGAAGATGTGGAC 660

Qy    661 cggcgcaatgaaaccctt[redacted]gacgccctgctcaccctctacccg 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661 CGGCGCAATGAAACCCTTCGACGGCAGCACCGGCCCCCGGCCCTGCTCACCCTCTACCCG 720

Qy    721 gcacctgacgaggatgaagccgtggaacgctgtagccgcccagagccaccccgcgagcac 780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    721 GCACCTGACGAGGATGAAGCCGTGGAACGCTGTAGCCGCCCAGAGCCACCCCGCGAGCAC 780

Qy    781 tttggacaaaggatcttggtcaagtgtctgtcgctcaagttcgagattgaaattgagccc 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    781 TTTGGACAAAGGATCTTGGTCAAGTGTCTGTCGCTCAAGTTCGAGATTGAAATTGAGCCC 840

Qy    841 atctttgggatcttggctctgtatgatgtgcgggagaaaaagaagatctcggagaacttc 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    841 ATCTTTGGGATCTTGGCTCTGTATGATGTGCGGGAGAAAAAGAAGATCTCGGAGAACTTC 900

Qy    901 tacttcgacctgaactcggactccatgaaggggctgcttcgggctcatggcacccaccct 960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    901 TACTTCGACCTGAACTCGGACTCCATGAAGGGGCTGCTTCGGGCTCATGGCACCCACCCT 960

Qy    961 gccatctccaccctggcccgctctgccatcttctctgtgacctaccctcacctgacatc 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    961 GCCATCTCCACCCTGGCCCGCTCTGCCATCTTCTCTGTGACCTACCCCTCACCTGACATC 1020

Qy   1021 ttcctggtcatcaagttggagaaggtgcttcagcaaggggacatcagtgagtgctgtgag 1080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1021 TTCCTGGTCATCAAGTTGGAGAAGGTGCTTCAGCAAGGGGACATCAGTGAGTGCTGTGAG 1080

Qy   1081 ccttacatggtgttgaaagaagtggacacagccaagaacaaagagaagctagagaagctg 1140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1081 CCTTACATGGTGTTGAAAGAAGTGGACACAGCCAAGAACAAAGAGAAGCTAGAGAAGCTG 1140

     1141 cgcctggcggccgagcagttctgcacccgcctgggccgctaccgcatgcccttcgcctgg 1200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1141 CGCCTGGCGGCCGAGCAGTTCTGCACCCGCCTGGGCCGCTACCGCATGCCCTTCGCCTGG 1200

Qy   1201 acggccgtgcacttggccaacatcgtgagcagcgctgggcagctggaccgggactctgac 1260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1201 ACGGCCGTGCACTTGGCCAACATCGTGAGCAGCGCTGGGCAGCTGGACCGGGACTCTGAC 1260

Qy   1261 tcggagggcgagcgccggccagcctggacagaccgccgccgtcgggggccccaggaccgg 1320
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1261 TCGGAGGGCGAGCGCCGGCCAGCCTGGACAGACCGCCGCCGTCGGGGGCCCCAGGACCGG 1320

Qy   1321 gcgagtagtggggacgacgcctgcagcttctctggcttccgtccagccacgctaactgtc 1380
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1321 GCGAGTAGTGGGGACGACGCCTGCAGCTTCTCTGGCTTCCGTCCAGCCACGCTAACTGTC 1380

Qy   1381 acaaacttctttaagcaggaggctgagcgactcagtgacgaggacctcttcaagttcctg 1440
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1381 ACAAACTTCTTTAAGCAGGAGGCTGAGCGACTCAGTGACGAGGACCTCTTCAAGTTCCTG 1440

     1441 gctgacatgaggcgcccgtcgtccctgctgcggcgactacgtcctgtgactgcccagctc 1500
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
     1441 GCTGACATGAGGCGCCCGTCGTCCCTGCTGCGGCGACTACGTCCTGTGACTGCCCAGCTC 1500

          [redacted]gatcgacatttctccggctcctgaaaatccccacttctgcctctccctgagctgctt 1560

Db   1501 AAGATCGACATTTCTCCGGCTCCTGAAAATCCCCACTTCTGCCTCTCCCCTGAGCTGCTT 1560

Qy   1561 catatcaagccctacccggaccccagggggccggcccaccaaggagattctggagttcccc 1620
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1561 CATATCAAGCCCTACCCGGACCCCAGGGGCCGGCCCACCAAGGAGATTCTGGAGTTCCCC 1620

Qy   1621 gcccgcgaagtctatgccccccataccagctacaggaacctgctgtacgtgtacccgcac 1680
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1621 GCCCGCGAAGTCTATGCCCCCCATACCAGCTACAGGAACCTGCTGTACGTGTACCCGCAC 1680

Qy   1681 agcctcaacttcagcagccgccagggctccgtgcgcaaccttgctgtgcgagtgcagtac 1740
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1681 AGCCTCAACTTCAGCAGCCGCCAGGGCTCCGTGCGCAACCTTGCTGTGCGAGTGCAGTAC 1740

Qy   1741 atgacaggcgaggaccccagccaggctctgccggtcatctttggcaagtccagctgcagt 1800
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1741 ATGACAGGCGAGGACCCCAGCCAGGCTCTGCCGGTCATCTTTGGCAAGTCCAGCTGCAGT 1800

Qy   1801 gaatttacccgcgaggccttcacaccggtggtctaccataacaagtccccccgagttctac 1860
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1801 GAATTTACCCGCGAGGCCTTCACACCGGTGGTCTACCATAACAAGTCCCCCGAGTTCTAC 1860

Qy   1861 gaggagttcaagctgcatcttccagcctgcgtgacagagaaccatcacctgctgttcacc 1920
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1861 GAGGAGTTCAAGCTGCATCTTCCAGCCTGCGTGACAGAGAACCATCACCTGCTGTTCACC 1920

Qy   1921 ttctaccatgtcagctgccagccccggccgggcactgccctggagacacccgtgggcttt 1980
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1921 TTCTACCATGTCAGCTGCCAGCCCCGGCCGGGCACTGCCCTGGAGACACCCGTGGGCTTT 1980

Qy   1981 acttggatcccactgctgcagcacgggcgcctgaggaccggccccttctgtctcccagtg 2040
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1981 ACTTGGATCCCACTGCTGCAGCACAGGCGCCTGAGGACCGGCCCCTTCTGTCTCCCAGTG 2040

Qy   2041 tctgtggaccagccgccgcccagctattccgtgctcacacccgatgtggcgcttccgggc 2100
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2041 TCTGTGGACCAGCCGCCGCCCAGCTATTCCGTGCTCACACCCGATGTGGCGCTTCCGGGC 2100

Qy   2101 atgcgctgggtggacggtcacaagggcgtgttcagtgtggagctcacagccgtgtcctct 2160
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2101 ATGCGCTGGGTGGACGGTCACAAGGGCGTGTTCAGTGTGGAGCTCACGGCCGTGTCCTCT 2160

Qy   2161 gtgcaccccccaggaccccctacctggacaaattcttcaccctggtgcacgtcctggaggag 2220
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2161 GTGCACCCCCAGGACCCCTACCTGGACAAATTCTTCACCCTGGTGCACGTCCTGGAGGAG 2220

Qy   2221 ggagccttcccattccggctcaaggacactgtgctgagcgagggcaacgtggagcaggag 2280
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2221 GGAGCCTTCCCATTCCGGCTCAAGGACACTGTGCTGAGCGAGGGCAACGTGGAGCAGGAG 2280

Qy   2281 ctgcgggccagtcttgcagcactgcgcctggccagccccgaaccccttgtggccttctcc 2340
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2281 CTGCGGGCCAGTCTTGCAGCACTGCGCCTGGCCAGCCCCGAACCCCTTGTGGCCTTCTCC 2340

Qy   2341 caccacgtgctggacaagctcgtgcgtctggtcatcaggccccccgatcatcagtggccag 2400
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2341 CACCACGTGCTGGACAAGCTCGTGCGTCTGGTCATCAGGCCCCCGATCATCAGTGGCCAG 2400

Qy   2401 attgtgaacctgggccgtggagcctttgaagcaatggcccatgtagtcagccttgttcac 2460
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2401 ATTGTGAACCTGGGCCGTGGAGCCTTTGAAGCAATGGCCCATGTAGTCAGCCTTGTTCAC 2460

Qy   2461 cggagcctggaggcagcccaggatgcccgcggtcactgcccacagctggctgcctacgtc 2520
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2461 CGGAGCCTGGAGGCAGCCCAGGATGCCCGCGGTCACTGCCCACAGCTGGCTGCCTACGTC 2520

Qy   2521 cactacgcctttcgccttcctggcactgagcccagcctcccggatggggcccctccagtg 2580
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2521 CACTACGCCTTTCGCCTTCCTGGCACTGAGCCCAGCCTCCCGGATGGGGCCCCTCCAGTG 2580

Qy   2581 acagtgcaggctgccacactggcccgtggctctggtcgccccgcaagcctctacctggcg 2640
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db  2581 ACAGTGCAGGCTGCCACACTGGCCCGTGGCTCTGGTCGCCCCGCAAGCCTCTACCTGGCG 2640
Qy  2641 cgttccaagagcatcagcagcagcaaccctgacctcgccgtggcccctggctctgtggat 2700
Db  2641 CGTTCCAAGAGCATCAGCAGCAGCAACCCTGACCTCGCCGTGGCCCCTGGCTCTGTGGAT 2700
Qy  2701 gacgaggtttcccgcatcctggccagcaagctgcttcacgaggagctggctctgcagtgg 2760
Db  2701 GACGAGGTTTCCCGCATCCTGGCCAGCAAGCTGCTTCACGAGGAGCTGGCTCTGCAGTGG 2760
Qy  2761 gtggtcagcagcagtgccgtacgcgaggccatcctccagcacgcctggttcttcttccag 2820
Db  2761 GTGGTCAGCAGCAGTGCCGTACGCGAGGCCATCCTCCAGCACGCCTGGTTCTTCTTCCAG 2820
Qy  2821 ctcatggtgaagagtatggcgctgcacctgctgcttggccagcgactagacacaccccgc 2880
Db  2821 CTCATGGTGAAGAGTATGGCGCTGCACCTGCTGCTTGGCCAGCGACTAGACACACCCCGC 2880
Qy  2881 aagctgcgcttccccggacgcttcctggacgacatcactgccttggtgggctctgtgggc 2940
Db  2881 AAGCTGCGCTTCCCCGGACGCTTCCTGGACGACATCACTGCCTTGGTGGGCTCTGTGGGC 2940
Qy  2941 ctggaggtcatcacccgtgtccacaaggatgtggagctggccgagcacctcaacgccagc 3000
Db  2941 CTGGAGGTCATCACCCGTGTCCACAAGGATGTGGAGCTGGCCGAGCACCTCAACGCCAGC 3000
Qy  3001 ctggctttcttcctcagtgaccttctgtccctggtggaccggggctttgtcttcagcctg 3060
Db  3001 CTGGCTTTCTTCCTCAGTGACCTTCTGTCCCTGGTGGACCGGGGCTTTGTCTTCAGCCTG 3060
Qy  3061 gtccgggcccactacaagcaggtggccacgcggctccagtcgtcccctaatccagcagcc 3120
Db  3061 GTCCGGGCCCACTACAAGCAGGTGGCCACGCGGCTCCAGTCGTCCCCTAATCCAGCAGCC 3120
Qy  3121 ctgctgaccctgcgcatggaattcacccgcatcctgtgcagccacgagcactacgtgacc 3180
Db  3121 CTGCTGACCCTGCGCATGGAATTCACCCGCATCCTGTGCAGCCACGAGCACTACGTGACC 3180
Qy  3181 ctcaacctcccctgctgcccccctgtcacctccagcctcgccctccccctctgtgtcctcc 3240
Db  3181 CTCAACCTCCCCTGCTGCCCCCTGTCACCTCCAGCCTCGCCCTCCCCCTCTGTGTCCTCC 3240
Qy  3241 accacctcccagagctccacattctccagccaagccccggaccccaaggtgaccagcatg 3300
Db  3241 ACCACCTCCCAGAGCTCCACCTTCTCCAGCCAAGCCCCGGACCCCAAGGTGACCAGCATG 3300
Qy  3301 ttcgaactgagtggaccattccggcagcagcacttcctagctgggctcctgctgacggag 3360
Db  3301 TTCGAACTGAGTGGACCATTCCGGCAGCAGCACTTCCTAGCTGGGCTCCTGCTGACGGAG 3360
Qy  3361 ctggcactggccctcgaaccgaggctgaaggggcattcctgttgcacaagaaggccatc 3420
Db  3361 CTGGCACTGGCCCTCGAACCGAGGCTGAAGGGGCATTCCTGTTGCACAAGAAGGCCATC 3420
Qy  3421 agtgctgtgcacagcctgctatgtggccatgacactgaccccgctacgccgaggccact 3480
Db  3421 AGTGCTGTGCACAGCCTGCTATGTGGCCATGACACTGACCCCCGCTACGCCGAGGCCACT 3480
Qy  3481 gtgaaggctcgtgtggccgagctgtacctgccactgctttcgattgcacgggataccttg 3540
Db  3481 GTGAAGGCTCGTGTGGCCGAGCTGTACCTGCCACTGCTTTCGATTGCACGGGATACCTTG 3540
Qy  3541 ccacggctgcatgactttgctgagggcccaggtcagcggtcaagactggcctcaatgctt 3600
Db  3541 CCACGGCTGCATGACTTTGCTGAGGGCCCAGGTCAGCGGTCAAGACTGGCCTCAATGCTT 3600
Qy  3601 gactcagacacagaaggcgaagggacattgcgggtaccatcaacccctctgtggccatg 3660
Db  3601 GACTCAGACACAGAAGGCGAAGGGACATTGCGGGTACCATCAACCCCTCTGTGGCCATG 3660
Qy  3661 gccattgctggtggccccctagccccctggctcccgggccagcatctcccaggggccacca 3720
Db  3661 GCCATTGCTGGTGGCCCCCTAGCCCCTGGCTCCCGGGCCAGCATCTCCCAGGGGCCACCA 3720

Qy  3721 acggcttctcgcgcaggctgtgccctctctgctgagtcaagccggaccttgctggcgtgt 3780
Db  3721 ACGGCTTCTCGCGCAGGCTGTGCCCTCTCTGCTGAGTCAAGCCGGACCTTGCTGGCGTGT 3780
Qy  3781 gtgctgtgggtgctgaaaaacaccgagccggcgctcctgcagcgctgggccactgacctg 3840
Db  3781 GTGCTGTGGGTGCTGAAAAACACCGAGCCGGCGCTCCTGCAGCGCTGGGCCACTGACCTG 3840
Qy  3841 acactcccccagctgggacgtctgttggacttgctgtacctttgcctggctgcctttgag 3900
Db  3841 ACACTCCCCCAGCTGGGACGTCTGTTGGACTTGCTGTACCTTTGCCTGGCTGCCTTTGAG 3900
Qy  3901 tacaaggggaaaaaggcctttgaacgcatcaacagcctcacattcaaaaaatctctggat 3960
Db  3901 TACAAGGGGAAAAAGGCCTTTGAACGCATCAACAGCCTCACATTCAAAAAATCTCTGGAT 3960
Qy  3961 atgaaggcgcggctagaggaagccattctgggtaccatcggagctcgacaagaaatggtt 4020
Db  3961 ATGAAGGCGCGGCTAGAGGAAGCCATTCTGGGTACCATCGGAGCTCGACAAGAAATGGTT 4020
Qy  4021 cggcgaagtcgtgagaggagcccgtttgggaatccagagaacgtgcgctggcggaagagc 4080
Db  4021 CGGCGAAGTCGTGAGAGGAGCCCGTTTGGGAATCCAGAGAACGTGCGCTGGCGGAAGAGC 4080
Qy  4081 gtcacacactggaagcaaacctcagaccgcgtggacaagaccaaggatgaaatggaacac 4140
Db  4081 GTCACACACTGGAAGCAAACCTCAGACCGCGTGGACAAGACCAAGGATGAAATGGAACAC 4140
Qy  4141 gaggccttggtggaagggaacctggcaaccgaggcaagcctagtggttctggacacactg 4200
Db  4141 GAGGCCTTGGTGGAAGGGAACCTGGCAACCGAGGCAAGCCTAGTGGTTCTGGACACACTG 4200
Qy  4201 gagatcatcgtgcagacggtgatgctttcagaagcccgggagagcgtcttgggggcagtg 4260
Db  4201 GAGATCATCGTGCAGACGGTGATGCTTTCAGAAGCCCGGGAGAGCGTCTTGGGGGCAGTG 4260
Qy  4261 ctgaaggttgtgctgtacagcctgggcagtgcccagagtgccctcttcttgcagcatggc 4320
Db  4261 CTGAAGGTTGTGCTGTACAGCCTGGGCAGTGCCCAGAGTGCCCTCTTCTTGCAGCATGGC 4320
Qy  4321 ctggccacccagagggcccttgtgtccaagttcccggagctgctgttcgaggaggacacg 4380
Db  4321 CTGGCCACCCAGAGGGCCCTTGTGTCCAAGTTCCCGGAGCTGCTGTTCGAGGAGGACACG 4380
Qy  4381 gagctgtgtgccgacctgtgcctgaggctcctacgacactgtggcagccgcatcagcacc 4440
Db  4381 GAGCTGTGTGCCGACCTGTGCCTGAGGCTCCTACGACACTGTGGCAGCCGCATCAGCACC 4440
Qy  4441 atccgcacgcacgccagcgcctcgctgtacctgctcatgcgacagaacttcgagatcggc 4500
Db  4441 ATCCGCACGCACGCCAGCGCCTCGCTGTACCTGCTCATGCGACAGAACTTCGAGATCGGC 4500
Qy  4501 cacaactttgcccgtgtgaagatgcaggtcaccatgtctctctcgtccctggtggggacg 4560
Db  4501 CACAACTTTGCCCGTGTGAAGATGCAGGTCACCATGTCTCTCTCGTCCCTGGTGGGGACG 4560
Qy  4561 acgcagaacttcagtgaagagcacctgcgacgttcactcaaaaccatcctcacctatgct 4620
Db  4561 ACGCAGAACTTCAGTGAAGAGCACCTGCGACGTTCACTCAAAACCATCCTCACCTATGCT 4620
Qy  4621 gaggaggacatggggctgcgggacagcaccttcgcagagcaggtccaggacctgatgttc 4680
Db  4621 GAGGAGGACATGGGGCTGCGGGACAGCACCTTCGCAGAGCAGGTCCAGGACCTGATGTTC 4680
Qy  4681 aacctgcacatgatcctgacggacacggtgaagatgaaggaacaccaggaggaccctgag 4740
Db  4681 AACCTGCACATGATCCTGACGGACACGGTGAAGATGAAGGAACACCAGGAGGACCCTGAG 4740
Qy  4741 atgctcatcgacctcatgtacagaattgcccggggctaccagggctcaccggaccttcgg 4800
Db  4741 ATGCTCATCGACCTCATGTACAGAATTGCCCGGGGCTACCAGGGCTCACCGGACCTTCGG 4800
```

A cont.

```
Qy    4801 ctgacctggttgcagaacatggccgggaagcacgcggagctgggcaaccacgccgaggcc 4860
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4801 CTGACCTGGTTGCAGAACATGGCCGGGAAGCACGCGGAGCTGGGCAACCACGCCGAGGCC 4860

Qy    4861 gcccagtgcatggtgcacgcggccgc 4886
           ||||||||||||||||||||||||||
Db    4861 GCCCAGTGCATGGTGCACGCGGCCGC 4886
```

RESULT 3
BC008335
-

Appendix B


```
seq_name: gb_pr:AB037816

seq_documentation_block:
LOCUS       AB037816                4886 bp    mRNA    linear   PRI 14-MAR-2000
DEFINITION  Homo sapiens mRNA for KIAA1395 protein, partial cds.
ACCESSION   AB037816
VERSION     AB037816.1  GI:7243170
KEYWORDS    .
SOURCE      Homo sapiens brain cDNA to mRNA, clone_lib:pBluescriptII SK plus
            clone:hj07927.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (sites)
  AUTHORS   Nagase,T., Kikuno,R., Ishikawa,K.I., Hirosawa,M. and Ohara,O.
  TITLE     Prediction of the coding sequences of unidentified human genes.
            XVI. The complete sequences of 150 new cDNA clones from brain which
            code for large proteins in vitro
  JOURNAL   DNA Res. 7 (1), 65-73 (2000)
  MEDLINE   20181126
REFERENCE   2  (bases 1 to 4886)
  AUTHORS   Ohara,O., Nagase,T. and Kikuno,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (31-JAN-2000) Osamu Ohara, Kazusa DNA Research Institute,
            Laboratory of DNA Technology; 1532-3 Yana, Kisarazu, Chiba
            292-0812, Japan (E-mail:cdnainfo@kazusa.or.jp,
            URL:http://www.kazusa.or.jp/huge/, Tel:+81-438-52-3913,
            Fax:+81-438-52-3914)
FEATURES             Location/Qualifiers
     source          1. .4886
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="hj07927"
                     /tissue_type="brain"
                     /clone_lib="pBluescriptII SK plus"
     gene            1. .4884
                     /gene="KIAA1395"
     CDS             <1. .>4884
                     /gene="KIAA1395"
                     /note="Start codon is not identified."
                     /codon_start=1
                     /product="KIAA1395 protein"
                     /protein_id="BAA92633.1"
                     /db_xref="GI:7243171"
                     /translation="DARTMAASERRAFAHKINRTVAAEVRKQVSRERSGSPHSSRRCS
                     SSLGVPLTEVVEPLDFEDVLLSRPPDAEPGPLRDLVEFPADDLELLLQPRECRTTEPG
                     IPKDEKLDAQVRAAVEMYIEDWVIVHRRYQYLSAAYSPVTTDTQRERQKGLPRQVFEQ
                     DASGDERSGPEDSNDSRRGSGSPEDTPRSSGASSIFDLRNLAADSLLPSLLERAAPED
                     VDRRNETLRRQHRPPALLTLYPAPDEDEAVERCSRPEPPREHEFGQRILVKCLSLKFEI
                     EIEPIFGILALYDVREKKKISENFYFDLNSDSMKGLLRAHGTHPAISTLARSAIFSVT
                     YPSPDIFLVIKLEKVLQQGDISECCEPYMVLKEVDTAKNKEKLEKLRLAAEQFCTRLG
                     RYRMPFAWTAVHLANIVSSAGQLDRDSDSEGERRPAWTDRRRRGPQDRASSGDDACSF
                     SGFRPATLTVTNFFKQEAERLSDEDLFKFLADMRRPSSLLRRLRPVTAQLKIDISPAP
                     ENPHFCLSPELLHIKPYPDPRGRPTKEILEFPAREVYAPHTSYRNLLYVYPHSLNFSS
                     RQGSVRNLAVRVQYMTGEDPSQALPVIFGKSSCSEFTREAFTPVVYHNKSPEFYEEFK
                     LHLPACVTENHHLLFTFYHVSCQPRPGTALETPVGFTWIPLLQHRRLRTGPFCLPVSV
                     DQPPPSYSVLTPDVALPGMRWVDGHKGVFSVELTAVSSVHPQDPYLDKFFTLVHVLEE
                     GAFPFRLKDTVLSEGNVEQELRASLAALRLASPEPLVAFSHHVLDKLVRLVIRPPIIS
                     GQIVNLGRGAFEAMAHVVSLVHRSLEAAQDARGHCPQLAAYVHYAFRLPGTEPSLPDG
                     APPVTVQAATLARGSGRPASLYLARSKSISSSNPDLAVAPGSVDDEVSRILASKLLHE
                     ELALQWVVSSSAVREAILQHAWFFFQLMVKSMALHLLLGQRLDTPRKLRFPGRFLDDI
                     TALVGSVGLEVITRVHKDVELAEHLNASLAFFLSDLLSLVDRGFVFSLVRAHYKQVAT
                     RLQSSPNPAALLTLRMEFTRILCSHEHYVTLNLPCCPLSPPASPSPSVSSTTSQSSTF
                     SSQAPDPKVTSMFELSGPFRQQHFLAGLLLTELALALEPEAEGAFLLHKKAISAVHSL
                     LCGHDTDPRYAEATVKARVAELYLPLLSIARDTLPRLHDFAEGPGQRSRLASMLDSDT
                     EGEGDIAGTINPSVAMAIAGGPLAPGSRASISQGPPTASRAGCALSAESSRTLLACVL
                     WVLKNTEPALLQRWATDLTLPQLGRLLDLLYLCLAAFEYKGKKAFERINSLTFKKSLD
                     MKARLEEAILGTIGARQEMVRRSRERSPFGNPENVRWRKSVTHWKQTSDRVDKTKDEM
                     EHEALVEGNLATEASLVVLDTLEIIVQTVMLSEARESVLGAVLKVVLYSLGSAQSALF
                     LQHGLATQRALVSKFPELLFEEDTELCADLCLRLLRHCGSRISTIRTHASASLYLLMR
                     QNFEIGHNFARVKMQVTMSLSSLVGTTQNFSEEHLRRSLKTILTYAEEDMGLRDSTFA
                     EQVQDLMFNLHMILTDTVKMKEHQEDPEMLIDLMYRIARGYQGSPDLRLTWLQNMAGK
                     HAELGNHAEAAQCMVHAA"
BASE COUNT      941 a   1605 c   1425 g    915 t
ORIGIN


alignment_scores:
                Quality: 8334.00              Length:   1624
                  Ratio:  5.135                 Gaps:      0
 Percent Similarity:  99.938   Percent Identity: 99.938
```

```
alignment_block:
 US-09-736-968A-2 x AB037816   ..

 Align seg 1/1  to: AB037816  from: 1  to: 4886

       1 MetAlaAlaSerGluArgArgAlaPheAlaHisLysIleAsnArgThrVa 17
         ||||||||||||||||||||||||||||||||||||||||||||||||||
      13 ATGGCTGCCTCCGAGCGCCGCGCCTTCGCGCACAAGATCAACAGGACGGT 62

      17 lAlaAlaGluValArgLysGlnValSerArgGluArgSerGlySerProH 34
         ||||||||||||||||||||||||||||||||||||||||||||||||||
      63 GGCCGCAGAGGTGCGGAAGCAGGTGTCCCGGGAACGCAGTGGCTCCCCCC 112

      34 isSerSerArgArgCysSerSerSerLeuGlyValProLeuThrGluVal 50
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     113 ACTCCAGCAGGCGCTGCAGCAGCTCCCTGGGGGTCCCACTGACTGAAGTT 162

      51 ValGluProLeuAspPheGluAspValLeuLeuSerArgProProAspAl 67
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     163 GTCGAGCCCCTGGACTTTGAGGATGTACTTCTGAGCCGGCCACCAGATGC 212

      67 aGluProGlyProLeuArgAspLeuValGluPheProAlaAspAspLeuG 84
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     213 TGAGCCCGGGCCCCTCAGGGACCTGGTAGAATTCCCAGCTGATGACTTGG 262

      84 luLeuLeuLeuGlnProArgGluCysArgThrThrGluProGlyIlePro 100
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     263 AGCTGCTGCTGCAGCCCCGGGAATGCCGGACCACGGAGCCCGGGATCCCC 312

     101 LysAspGluLysLeuAspAlaGlnValArgAlaAlaValGluMetTyrIl 117
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     313 AAGGATGAAAAACTGGATGCCCAGGTGAGGGCCGCGGTGGAGATGTATAT 362

     117 eGluAspTrpValIleValHisArgArgTyrGlnTyrLeuSerAlaAlaT 134
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     363 TGAGGACTGGGTCATTGTCCACAGAAGGTATCAGTACCTGAGTGCAGCAT 412

     134 yrSerProValThrThrAspThrGlnArgGluArgGlnLysGlyLeuPro 150
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     413 ACAGCCCCGTCACCACAGACACACAGCGGGAGCGACAGAAGGGCCTCCCC 462

     151 ArgGlnValPheGluGlnAspAlaSerGlyAspGluArgSerGlyProGl 167
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     463 CGCCAGGTCTTTGAGCAGGATGCTTCTGGAGACGAGAGGTCCGGCCCTGA 512

     167 uAspSerAsnAspSerArgArgGlySerGlySerProGluAspThrProA 184
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     513 GGACTCGAATGACTCCCGGCGTGGCTCGGGCTCCCCGGAAGACACCCCTC 562

     184 rgSerSerGlyAlaSerSerIlePheAspLeuArgAsnLeuAlaAlaAsp 200
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     563 GAAGCAGTGGTGCCTCTAGCATCTTCGACCTGAGGAACCTGGCAGCTGAC 612

     201 SerLeuLeuProSerLeuLeuGluArgAlaAlaProGluAspValAspAr 217
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     613 TCATTGCTGCCCTCTCTGCTAGAGCGGGCGGCCCCAGAAGATGTGGACCG 662

     217 gArgAsnGluThrLeuArgArgGlnHisArgProProAlaLeuLeuThrL 234
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     663 GCGCAATGAAACCCTTCGACGGCAGCACCGGCCCCCGGCCCTGCTCACCC 712

     234 euTyrProAlaProAspGluAspGluAlaValGluArgCysSerArgPro 250
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     713 TCTACCCGGCACCTGACGAGGATGAAGCCGTGGAACGCTGTAGCCGCCCA 762

     251 GluProProArgGluHisPheGlyGlnArgIleLeuValLysCysLeuSe 267
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     763 GAGCCACCCCGCGAGCACTTTGGACAAAGGATCTTGGTCAAGTGTCTGTC 812

     267 rLeuLysPheGluIleGluIleGluProIlePheGlyIleLeuAlaLeuT 284
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     813 GCTCAAGTTCGAGATTGAAATTGAGCCCATCTTTGGGATCTTGGCTCTGT 862

     284 yrAspValArgGluLysLysLysIleSerGluAsnPheTyrPheAspLeu 300
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     863 ATGATGTGCGGGAGAAAAAGAAGATCTCGGAGAACTTCTACTTCGACCTG 912

     301 AsnSerAspSerMetLysGlyLeuLeuArgAlaHisGlyThrHisProAl 317
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     913 AACTCGGACTCCATGAAGGGGCTGCTTCGGGCTCATGGCACCCACCCTGC 962

     317 aIleSerThrLeuAlaArgSerAlaIlePheSerValThrTyrProSerP 334
         ||||||||||||||||||||||||||||||||||||||||||||||||||
     963 CATCTCCACCCTGGCCCGCTCTGCCATCTTCTCTGTGACCTACCCCTCAC 1012

     334 roAspIlePheLeuValIleLysLeuGluLysValLeuGlnGlnGlyAsp 350
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    1013 CTGACATCTTCCTGGTCATCAAGTTGGAGAAGGTGCTTCAGCAAGGGGAC 1062

     351 IleSerGluCysCysGluProTyrMetValLeuLysGluValAspThrAl 367
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    1063 ATCAGTGAGTGCTGTGAGCCTTACATGGTGTTGAAAGAAGTGGACACAGC 1112

     367 aLysAsnLysGluLysLeuGluLysLeuArgLeuAlaAlaGluGlnPheC 384
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    1113 CAAGAACAAAGAGAAGCTAGAGAAGCTGCGCCTGGCGGCCGAGCAGTTCT 1162

     384 ysThrArgLeuGlyArgTyrArgMetProPheAlaTrpThrAlaValHis 400
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    1163 GCACCCGCCTGGGCCGCTACCGCATGCCCTTCGCCTGGACGGCCGTGCAC 1212

     401 LeuAlaAsnIleValSerSerAlaGlyGlnLeuAspArgAspSerAspSe 417
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    1213 TTGGCCAACATCGTGAGCAGCGCTGGGCAGCTGGACCGGGACTCTGACTC 1262

     417 rGluGlyGluArgArgProAlaTrpThrAspArgArgArgArgGlyProG 434
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    1263 GGAGGGCGAGCGCCGGCCAGCCTGGACAGACCGCCGCCGTCGGGGGCCCC 1312

     434 lnAspArgAlaSerSerGlyAspAspAlaCysSerPheSerGlyPheArg 450
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    1313 AGGACCGGGCGAGTAGTGGGGACGACGCCTGCAGCTTCTCTGGCTTCCGT 1362

     451 ProAlaThrLeuThrValThrAsnPhePheLysGlnGluAlaGluArgLe 467
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    1363 CCAGCCACGCTAACTGTCACAAACTTCTTTAAGCAGGAGGCTGAGCGACT 1412

     467 uSerAspGluAspLeuPheLysPheLeuAlaAspMetArgArgProSerS 484
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    1413 CAGTGACGAGGACCTCTTCAAGTTCCTGGCTGACATGAGGCGCCCGTCGT 1462

     484 erLeuLeuArgArgLeuArgProValThrAlaGlnLeuLysIleAspIle 500
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    1463 CCCTGCTGCGGCGACTACGTCCTGTGACTGCCCAGCTCAAGATCGACATT 1512

     501 SerProAlaProGluAsnProHisPheCysLeuSerProGluLeuLeuHi 517
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    1513 TCTCCGGCTCCTGAAAATCCCCACTTCTGCCTCTCCCCTGAGCTGCTTCA 1562

     517 sIleLysProTyrProAspProArgGlyArgProThrLysGluIleLeuG 534
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    1563 TATCAAGCCCTACCCGGACCCCAGGGGCCGGCCCACCAAGGAGATTCTGG 1612

     534 luProAlaArgGluValTyrAlaProHisThrSerTyrArgAsnLeu 550
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    1613 AGTTCCCCGCCCGCGAAGTCTATGCCCCCCATACCAGCTACAGGAACCTG 1662

     551 LeuTyrValTyrProHisSerLeuAsnPheSerSerArgGlnGlySerVa 567
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    1663 CTGTACGTGTACCCGCACAGCCTCAACTTCAGCAGCCGCCAGGGCTCCGT 1712

     567 lArgAsnLeuAlaValArgValGlnTyrMetThrGlyGluAspProSerG 584
         ||||||||||||||||||||||||||||||||||||||||||||||||||
    1713 GCGCAACCTTGCTGTGCGAGTGCAGTACATGACAGGCGAGGACCCCAGCC 1762

     584 lnAlaLeuProValIlePheGlyLysSerSerCysSerGluPheThrArg 600
```

```
1763 AGGCTCTGCCGGTCATCTTTGGCAAGTCCAGCTGCAGTGAATTTACCCGC 1812

 601 GluAlaPheThrProValValTyrHisAsnLysSerProGluPheTyrGl 617
1813 GAGGCCTTCACACCGGTGGTCTACCATAACAAGTCCCCCGAGTTCTACGA 1862

 617 uGluPheLysLeuHisLeuProAlaCysValThrGluAsnHisHisLeuL 634
1863 GGAGTTCAAGCTGCATCTTCCAGCCTGCGTGACAGAGAACCATCACCTGC 1912

 634 euPheThrPheTyrHisValSerCysGlnProArgProGlyThrAlaLeu 650
1913 TGTTCACCTTCTACCATGTCAGCTGCCAGCCCCGGCCGGGCACTGCCCTG 1962

 651 GluThrProValGlyPheThrTrpIleProLeuLeuGlnHisGlyArgLe 667
1963 GAGACACCCGTGGGCTTTACTTGGATCCCACTGCTGCAGCACAGGCGCCT 2012

 667 uArgThrGlyProPheCysLeuProValSerValAspGlnProProProS 684
2013 GAGGACCGGCCCCTTCTGTCTCCCAGTGTCTGTGGACCAGCCGCCGCCCA 2062

 684 erTyrSerValLeuThrProAspValAlaLeuProGlyMetArgTrpVal 700
2063 GCTATTCCGTGCTCACACCCGATGTGGCGCTTCCGGGCATGCGCTGGGTG 2112

 701 AspGlyHisLysGlyValPheSerValGluLeuThrAlaValSerServa 717
2113 GACGGTCACAAGGGCGTGTTCAGTGTGGAGCTCACGGCCGTGTCCTCTGT 2162

 717 lHisProGlnAspProTyrLeuAspLysPhePheThrLeuValHisValL 734
2163 GCACCCCCAGGACCCCTACCTGGACAAATTCTTCACCCTGGTGCACGTCC 2212

 734 euGluGluGlyAlaPheProPheArgLeuLysAspThrValLeuSerGlu 750
2213 TGGAGGAGGGAGCCTTCCCATTCCGGCTCAAGGACACTGTGCTGAGCGAG 2262

 751 GlyAsnValGluGlnGluLeuArgAlaSerLeuAlaAlaLeuArgLeuAl 767
2263 GGCAACGTGGAGCAGGAGCTGCGGGCCAGTCTTGCAGCACTGCGCCTGGC 2312

 767 aSerProGluProLeuValAlaPheSerHisHisValLeuAspLysLeuV 784
2313 CAGCCCCGAACCCCTTGTGGCCTTCTCCCACCACGTGCTGGACAAGCTCG 2362

 784 alArgLeuValIleArgProProIleIleSerGlyGlnIleValAsnLeu 800
2363 TGCGTCTGGTCATCAGGCCCCCGATCATCAGTGGCCAGATTGTGAACCTG 2412

 801 GlyArgGlyAlaPheGluAlaMetAlaHisValValSerLeuValHisAr 817
2413 GGCCGTGGAGCCTTTGAAGCAATGGCCCATGTAGTCAGCCTTGTTCACCG 2462

 817 gSerLeuGluAlaAlaGlnAspAlaArgGlyHisCysProGlnLeuAlaA 834
2463 GAGCCTGGAGGCAGCCCAGGATGCCCGCGGTCACTGCCCACAGCTGGCTG 2512

 834 laTyrValHisTyrAlaPheArgLeuProGlyThrGluProSerLeuPro 850
2513 CCTACGTCCACTACGCCTTTCGCCTTCCTGGCACTGAGCCCAGCCTCCCG 2562

 851 AspGlyAlaProProValThrValGlnAlaAlaThrLeuAlaArgGlySe 867
2563 GATGGGGCCCCTCCAGTGACAGTGCAGGCTGCCACACTGGCCCGTGGCTC 2612

 867 rGlyArgProAlaSerLeuTyrLeuAlaArgSerLysSerIleSerSerS 884
2613 TGGTCGCCCCGCAAGCCTCTACCTGGCGCGTTCCAAGAGCATCAGCAGCA 2662

 884 erAsnProAspLeuAlaValAlaProGlySerValAspAspGluValSer 900
```

```
2663 GCAACCCTGACCTCGCCGTGGCCCCTGGCTCTGTGGATGACGAGGTTTCC 2712

 901 ArgIleLeuAlaSerLysLeuLeuHisGluGluLeuAlaLeuGlnTrpVa 917
2713 CGCATCCTGGCCAGCAAGCTGCTTCACGAGGAGCTGGCTCTGCAGTGGGT 2762

 917 lValSerSerSerAlaValArgGluAlaIleLeuGlnHisAlaTrpPheP 934
2763 GGTCAGCAGCAGTGCCGTACGCGAGGCCATCCTCCAGCACGCCTGGTTCT 2812

 934 hePheGlnLeuMetValLysSerMetAlaLeuHisLeuLeuLeuGlyGln 950
2813 TCTTCCAGCTCATGGTGAAGAGTATGGCGCTGCACCTGCTGCTTGGCCAG 2862

 951 ArgLeuAspThrProArgLysLeuArgPheProGlyArgPheLeuAspAs 967
2863 CGACTAGACACACCCCGCAAGCTGCGCTTCCCCGGACGCTTCCTGGACGA 2912

 967 pIleThrAlaLeuValGlySerValGlyLeuGluValIleThrArgValH 984
2913 CATCACTGCCTTGGTGGGCTCTGTGGGCCTGGAGGTCATCACCCGTGTCC 2962

 984 isLysAspValGluLeuAlaGluHisLeuAsnAlaSerLeuAlaPhePhe 1000
2963 ACAAGGATGTGGAGCTGGCCGAGCACCTCAACGCCAGCCTGGCTTTCTTC 3012

1001 LeuSerAspLeuLeuSerLeuValAspArgGlyPheValPheSerLeuVa 1017
3013 CTCAGTGACCTTCTGTCCCTGGTGGACCGGGGCTTTGTCTTCAGCCTGGT 3062

1017 lArgAlaHisTyrLysGlnValAlaThrArgLeuGlnSerSerProAsnP 1034
3063 CCGGGCCCACTACAAGCAGGTGGCCACGCGGCTCCAGTCGTCCCCTAATC 3112

1034 roAlaAlaLeuLeuThrLeuArgMetGluPheThrArgIleLeuCysSer 1050
3113 CAGCAGCCCTGCTGACCCTGCGCATGGAATTCACCCGCATCCTGTGCAGC 3162

1051 HisGluHisTyrValThrLeuAsnLeuProCysCysProLeuSerProPr 1067
3163 CACGAGCACTACGTGACCCTCAACCTCCCCTGCTGCCCCCTGTCACCTCC 3212

1067 oAlaSerProSerProSerValSerSerThrThrSerGlnSerSerThrP 1084
3213 AGCCTCGCCCTCCCCCTCTGTGTCCTCCACCACCTCCCAGAGCTCCACCT 3262

1084 heSerSerGlnAlaProAspProLysValThrSerMetPheGluLeuSer 1100
3263 TCTCCAGCCAAGCCCCGGACCCCAAGGTGACCAGCATGTTCGAACTGAGT 3312

1101 GlyProPheArgGlnGlnHisPheLeuAlaGlyLeuLeuLeuThrGluLe 1117
3313 GGACCATTCCGGCAGCAGCACTTCCTAGCTGGGCTCCTGCTGACGGAGCT 3362

1117 uAlaLeuAlaLeuGluProGluAlaGluGlyAlaPheLeuLeuHisLysL 1134
3363 GGCACTGGCCCTCGAACCTGAGGCTGAAGGGGCATTCCTGTTGCACAAGA 3412

1134 ysAlaIleSerAlaValHisSerLeuLeuCysGlyHisAspThrAspPro 1150
3413 AGGCCATCAGTGCTGTGCACAGCCTGCTATGTGGCCATGACACTGACCCC 3462

1151 ArgTyrAlaGluAlaThrValLysAlaArgValAlaGluLeuTyrLeuPr 1167
3463 CGCTACGCCGAGGCCACTGTGAAGGCTCGTGTGGCCGAGCTGTACCTGCC 3512

1167 oLeuLeuSerIleAlaArgAspThrLeuProArgLeuHisAspPheAlaG 1184
3513 ACTGCTTTCGATTGCACGGGATACCTTGCCACGGCTGCATGACTTTGCTG 3562

1184 luGlyProGlyGlnArgSerArgLeuAlaSerMetLeuAspSerAspThr 1200
3563 AGGGCCCAGGTCAGCGGTCAAGACTGGCCTCAATGCTTGACTCAGACACA 3612
```

```
1201 GluGlyGluGlyAspIleAlaGlyThrIleAsnProSerValAlaMetAl 1217
     ||||||||||||||||||||||||||||||||||||||||||||||||||
3613 GAAGGCGAAGGGGACATTGCGGGTACCATCAACCCCTCTGTGGCCATGGC 3662

1217 aIleAlaGlyGlyPrcLeuAlaProGlySerArgAlaSerIleSerGlnG 1234
     ||||||||||||||||||||||||||||||||||||||||||||||||||
3663 CATTGCTGGTGGCCCCCTAGCCCCTGGCTCCCGGGCCAGCATCTCCCAGG 3712

1234 lyProProThrAlaSerArgAlaGlyCysAlaLeuSerAlaGluSerSer 1250
     ||||||||||||||||||||||||||||||||||||||||||||||||||
3713 GGCCACCAACGGCTTCTCGCGCAGGCTGTGCCCTCTCTGCTGAGTCAAGC 3762

1251 ArgThrLeuLeuAlaCysValLeuTrpValLeuLysAsnThrGluProAl 1267
     ||||||||||||||||||||||||||||||||||||||||||||||||||
3763 CGGACCTTGCTGGCGTGTGTGCTGTGGGTGCTGAAAAACACCGAGCCGGC 3812

1267 aLeuLeuGlnArgTrpAlaThrAspLeuThrLeuProGlnLeuGlyArgL 1284
     ||||||||||||||||||||||||||||||||||||||||||||||||||
3813 GCTCCTGCAGCGCTGGCCCACTGACCTGACACTCCCCCAGCTGGGACGTC 3862

1284 euLeuAspLeuLeuTyrLeuCysLeuAlaAlaPheGluTyrLysGlyLys 1300
     ||||||||||||||||||||||||||||||||||||||||||||||||||
3863 TGTTGGACTTGCTGTACCTTTGCCTGGCTGCCTTTGAGTACAAGGGGAAA 3912

1301 LysAlaPheGluArgIleAsnSerLeuThrPheLysLysSerLeuAspMe 1317
     ||||||||||||||||||||||||||||||||||||||||||||||||||
3913 AAGGCCTTTGAACGCATCAACAGCCTCACATTCAAAAAATCTCTGGATAT 3962

1317 tLysAlaArgLeuGluGluAlaIleLeuGlyThrIleGlyAlaArgGlnG 1334
     ||||||||||||||||||||||||||||||||||||||||||||||||||
3963 GAAGGCGCGGCTAGAGGAAGCCATTCTGGGTACCATCGGAGCTCGACAAG 4012

1334 luMetValArgArgSerArgGluArgSerProPheGlyAsnProGluAsn 1350
     ||||||||||||||||||||||||||||||||||||||||||||||||||
4013 AAATGGTTCGGCGAAGTCGTGAGAGGAGCCCGTTTGGGAATCCAGAGAAC 4062

1351 ValArgTrpArgLysSerValThrHisTrpLysGlnThrSerAspArgVa 1367
     ||||||||||||||||||||||||||||||||||||||||||||||||||
4063 GTGCGCTGGCGGAAGAGCGTCACACACTGGAAGCAAACCTCAGACCGCGT 4112

1367 lAspLysThrLysAspGlnMetGluHisGluAlaLeuValGluGlyAsnL 1384
     ||||||||||||||||||||||||||||||||||||||||||||||||||
4113 GGACAAGACCAAGGATGAAATGGAACACGAGGCCTTGGTGGAAGGGAACC 4162

1384 euAlaThrGluAlaSerLeuValValLeuAspThrLeuGluIleIleVal 1400
     ||||||||||||||||||||||||||||||||||||||||||||||||||
4163 TGGCAACCGAGGCAAGCCTAGTGGTTCTGGACACACTGGAGATCATCGTG 4212

1401 GlnThrValMetLeuSerGluAlaArgGluSerValLeuGlyAlaValLe 1417
     ||||||||||||||||||||||||||||||||||||||||||||||||||
4213 CAGACGGTGATGCTTTCAGAAGCCCGGGAGAGCGTCTTGGGGGCAGTGCT 4262

1417 uLysValValLeuTyrSerLeuGlySerAlaGlnSerAlaLeuPheLeuG 1434
     ||||||||||||||||||||||||||||||||||||||||||||||||||
4263 GAAGGTTGTGCTGTACAGCCTGGGCAGTGCCCAGAGTGCCCTCTTCTTGC 4312

1434 lnHisGlyLeuAlaThrGlnArgAlaLeuValSerLysProProGluLeu 1450
     ||||||||||||||||||||||||||||||||||||||||||||||||||
4313 AGCATGGCCTGGCCACCCAGAGGGCCCTTGTGTCCAAGTTCCCGGAGCTG 4362

1451 LeuPheGluGluAspThrGluLeuCysAlaAspLeuCysLeuArgLeuLe 1467
     ||||||||||||||||||||||||||||||||||||||||||||||||||
4363 CTGTTCGAGGAGGACACGGAGCTGTGTGCCGACCTGTGCCTGAGGCTCCT 4412

1467 uArgHisCysGlySerArgIleSerThrIleArgThrHisAlaSerAlaS 1484
     ||||||||||||||||||||||||||||||||||||||||||||||||||
4413 ACGACACTGTGGCAGCCGCATCAGCACCATCCGCACGCACGCCAGCGCCT 4462

    erLeuTyrLeuLeuMetArgGlnAsnPheGluIleGlyHisAsnPheAla 1500
    ||||||||||||||||||||||||||||||||||||||||||||||||||
    GCTGTACCTGCTCATGCGACAGAACTTCGAGATCGGCCACAACTTTGCC 4512

1501 ArgValLysMetGlnValThrMetSerLeuSerSerLeuValGlyThrTh 1517
     ||||||||||||||||||||||||||||||||||||||||||||||||||
4513 CGTGTGAAGATGCAGGTCACCATGTCTCTCTCGTCCCTGGTGGGGACGAC 4562

1517 rGlnAsnPheSerGluGluHisLeuArgArgSerLeuLysThrIleLeuT 1534
     ||||||||||||||||||||||||||||||||||||||||||||||||||
4563 GCAGAACTTCAGTGAAGAGCACCTGCGACGTTCACTCAAAACCATCCTCA 4612

1534 hrTyrAlaGluGluAspMetGlyLeuArgAspSerThrPheAlaGluGln 1550
     ||||||||||||||||||||||||||||||||||||||||||||||||||
4613 CCTATGCTGAGGAGGACATGGGGCTGCGGGACAGCACCTTCGCAGAGCAG 4662

1551 ValGlnAspLeuMetPheAsnLeuHisMetIleLeuThrAspThrValLy 1567
     ||||||||||||||||||||||||||||||||||||||||||||||||||
4663 GTCCAGGACCTGATGTTCAACCTGCACATGATCCTGACGGACACGGTGAA 4712

1567 sMetLysGluHisGlnGluAspProGluMetLeuIleAspLeuMetTyrA 1584
     ||||||||||||||||||||||||||||||||||||||||||||||||||
4713 GATGAAGGAACACCAGGAGGACCCTGAGATGCTCATCGACCTCATGTACA 4762

1584 rgIleAlaArgGlyTyrGlnGlySerProAspLeuArgLeuThrTrpLeu 1600
     ||||||||||||||||||||||||||||||||||||||||||||||||||
4763 GAATTGCCCGGGGCTACCAGGGCTCACCGGACCTTCGGCTGACCTGGTTG 4812

1601 GlnAsnMetAlaGlyLysHisAlaGluLeuGlyAsnHisAlaGluAlaAl 1617
     ||||||||||||||||||||||||||||||||||||||||||||||||||
4813 CAGAACATGGCCGGGAAGCACGCGGAGCTGGGCAACCACGCCGAGGCCGC 4862

1617 aGlnCysMetValHisAlaAla 1624
     ||||||||||||||||||||||
4863 CCAGTGCATGGTGCACGCGGCC 4884
```

Appendix C


```
RESULT 10
AI198543/c
LOCUS       AI198543                794 bp    mRNA    linear   EST 10-NOV-1998
DEFINITION  qf49f11.x1 Soares_testis_NHT Homo sapiens cDNA clone IMAGE:1753389
            3' similar to WP:F46H5.4 CE04590 RAT TRG GENE PRODUCT ;, mRNA
            sequence.
ACCESSION   AI198543
VERSION     AI198543.1  GI:3751149
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 794)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima Bonaldo
            , Ph.D.
            cDNA Library Arrayed by: Greg Lennon, Ph.D.
            DNA Sequencing by: Washington University Genome Sequencing Center
            Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html
            Insert Length: 856   Std Error: 0.00
            Seq primer: -40UP from Gibco
            High quality sequence stop: 444.
FEATURES             Location/Qualifiers
     source          1. .794
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:1753389"
                     /clone_lib="Soares_testis_NHT"
                     /sex="male"
                     /lab_host="DH10B"
                     /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
                     polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
                     was prepared from mRNA obtained from Clontech Laboratories
                     , Inc., and primed with  a Not I - oligo(dT) primer [5'
                     TGTTACCAATCTGAAGTGGGAGCGGCCGCCCAATTTTTTTTTTTTTTT 3'].
                     Double-stranded cDNA was ligated to Eco RI adaptors
                     (Pharmacia), digested with Not I and cloned into the Not I
                     and Eco RI sites of the modified pT7T3 vector. Library
                     went through one round of normalization to Cot5, and was
                     constructed  by Bento Soares and M. Fatima Bonaldo. "
BASE COUNT      149 a    219 c    236 g    190 t
ORIGIN
```

```
  Query Match            11.5%;  Score 735;  DB 9;  Length 794;
  Best Local Similarity  97.7%;  Pred. No. 1.9e-122;
  Matches  777;  Conservative    0;  Mismatches   15;  Indels    3;  Gaps    3;
```

```
Qy   5578 ttcacgccggatgggcgcgcacacggggagctgcccgagcaacacaagcgtaagacgctg 5637
          ||||||||| |||||||||||||||||| || |||||||||||| ||||| |||||||  |
Db    794 TTCACGCCGAATGGGCGCGCACACGGGAAGTTGCCCGAGCAACCCAAGCTTAAGACGTGG 735

Qy   5638 ctcagcaccgaccacgccttcccctacatcaagactcgcatccgtgtgtgccaccgggag 5697
          | |||| |||||||||||||||||||||||||||||||||||||||||| ||| ||||||
Db    734 TTAAGCA-CGACCACGCCTTCCCCTACATCAAGACTCGCATCCGTGTGTCCCA-CGGGAG 677

Qy   5698 gagacggtgctgacgccagtggaggtggccatcgaggacatgcagaagaagacacgggag 5757
          || ||||||| | |||||||||||||||||||||||||||||| ||||||||||||||||||
Db    676 GAAACGGTGTTAACGCCAGTGGAGGTGGCCATCGAGGACATCCAGAAGAAGACACGGGAG 617

Qy   5758 ctggcctttgccaccgagcaggacccaccagatgctaagatgctacagatggtgcttcag 5817
          ||||||||| |||||||||||||||||||||||||||||||||||||||||||||||||||
Db    616 TTGGCCTTT-CCACCGAGCAGGACCCACCAGATGCTAAGATGCTACAGATGGTGCTTCAG 558

Qy   5818 ggctctgtagggcccaccgtgaaccagggtcccctggaggtggcccaggtgtttttagca 5877
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    557 GGCTCTGTAGGGCCCACCGTGAACCAGGGTCCCCTGGAGGTGGCCCAGGTGTTTTTAGCA 498

Qy   5878 gagatcccggaagaccccaagctcttccggcatcacaacaaattgcggctctgcttcaag 5937
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    497 GAGATCCCGGAAGACCCCAAGCTCTTCCGGCATCACAACAAATTGCGGCTCTGCTTCAAG 438

Qy   5938 gacttctgcaagaaatgtgaggatgcgctgcggaaaaataaggccctgattgggccggac 5997
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    437 GACTTCTGCAAGAAATGTGAGGATGCGCTGCGGAAAAATAAGGCCCTGATTGGGCCGGAC 378

Qy   5998 cagaaggagtaccaccgtgagctggagcgcaactactgccgcctgcgggaggctctgcag 6057
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    377 CAGAAGGAGTACCACCGTGAGCTGGAGCGCAACTACTGCCGCCTGCGGGAGGCTCTGCAG 318

Qy   6058 cccctgcttacccagcgcctgccccagctgatggcacccaccccacccggcctcaggaac 6117
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    317 CCCCTGCTTACCCAGCGCCTGCCCCAGCTGATGGCACCCACCCCACCCGGCCTCAGGAAC 258

Qy   6118 tccttgaacagagcaagtttccgaaaggcagacctctgagcccacaaggaccaaagctgt 6177
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    257 TCCTTGAACAGAGCAAGTTTCCGAAAGGCAGACCTCTGAGCCCACAAGGACCAAAGCTGT 198

Qy   6178 acctagaggaaccagcacccgggcctcagctgtctgtgctgcgaggggagtctgccctgg 6237
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||H
Db    197 ACCTAGAGGAACCAGCACCCGGGCCTCAGCTGTCTGTGCTGCGAGGGGAGTCTGCCCTGG 138

Qy   6238 tgcccactgggctgtggggtgaccacactgtacttggggctgggccctctgcccctgtgt 6297
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    137 TGCCCACTGGGCTGTGGGGTGACCACACTGTACTTGGGGCTGGGCCCTCTGCCCCTGTGT 78

Qy   6298 ccccatctgtgtgcactgatgcttcctccctttttaatttaaaatggtttttataagca 6357
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     77 CCCCATCTGTGTGCACTGATGCTTCCTCCCTTTTTTAATTTAAAATGGTTTTTATAAGCA 18

Qy   6358 aaaaaaaaaaaaaaa 6372
          |||||||||||||||
Db    17 AAAAAAAAAAAAAAA 3
```